

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2002-542526

(P2002-542526A)

(43) 公表日 平成14年12月10日 (2002. 12. 10)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
G 0 6 F 12/00	5 3 1	G 0 6 F 12/00	5 3 1 D 5 B 0 1 8
	5 3 5		5 3 5 R 5 B 0 8 2
12/16	3 1 0	12/16	3 1 0 J 5 B 0 8 3
13/00	3 0 1	13/00	3 0 1 P

審査請求 有 予備審査請求 未請求(全 43 頁)

(21) 出願番号 特願2000-600177(P2000-600177)
 (86) (22) 出願日 平成11年10月12日 (1999. 10. 12)
 (85) 翻訳文提出日 平成12年10月17日 (2000. 10. 17)
 (86) 国際出願番号 P C T / U S 9 9 / 2 3 9 3 9
 (87) 国際公開番号 W O 0 0 / 4 9 5 0 0
 (87) 国際公開日 平成12年8月24日 (2000. 8. 24)
 (31) 優先権主張番号 0 9 / 2 5 1 , 8 1 2
 (32) 優先日 平成11年2月17日 (1999. 2. 17)
 (33) 優先権主張国 米国 (U S)
 (81) 指定国 E P (A T , B E , C H , C Y ,
 D E , D K , E S , F I , F R , G B , G R , I E , I
 T , L U , M C , N L , P T , S E) , I L , J P , K
 R

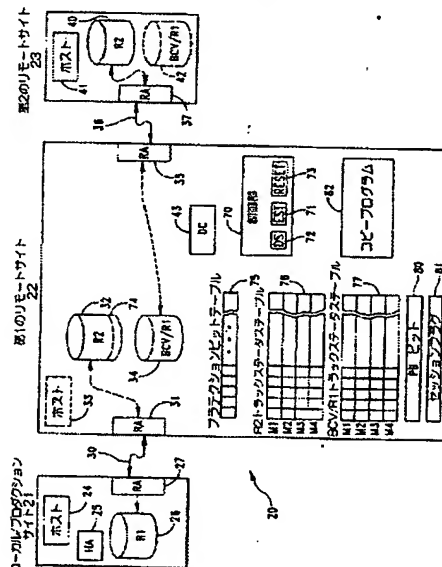
(71) 出願人 イーエムシー コーポレイション
 アメリカ合衆国 マサチューセッツ州
 01748 ホプキントン サウス ストリート
 171
 (72) 発明者 ガーニュ マチュール
 アメリカ合衆国 マサチューセッツ州
 02114 ポストン ジョイ ストリート
 43 アパートメント 5
 (72) 発明者 オフェク ユーヴァル
 アメリカ合衆国 マサチューセッツ州
 01748 フレミングラム ランターンズ ロ
 ード 20
 (74) 代理人 弁理士 中村 稔 (外9名)

最終頁に続く

(54) 【発明の名称】 冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

(57) 【要約】

プロダクションデータ処理サイト (21) のようなデータ変更装置からリモートデータ受信サイト (23) へデータを転送するためのデータ記憶設備 (22)。このデータ記憶設備は、データ変更装置によって生成されたデータ中のそれぞれの変更を記録するための第1のデータ記憶装置を含む。レジスタセット (75) は、それぞれの変更をトラック毎に記録する。第2のデータ記憶装置 (34) は、第1及び第2の動作モードを有する。第1の動作モードの間中に、第2のデータ記憶装置は、第1のデータ記憶装置のミラーになる。第2の動作モードの間中に、第2のデータ記憶装置は、ミラーとして働くことをやめ、データ受信サイトへのデータの転送のためのソースになる。変更された情報のみ、すなわち変更された特定のトラックが、第2の動作モード中の連続した動作の間中に転送される。ローカルのプロダクションサイトからのコマンドは、第1と第2の動作モードの間の転送を開始する。



【特許請求の範囲】

【請求項1】 データ処理システムにおける、データ変更手段とデータ受信手段の間に介在するデータ記憶設備であって、

- A. データを前記データ変更手段から受信するための、前記データ変更手段との接続のための第1のデータ記憶装置と、
- B. 第2のデータ記憶装置と、
- C. 前記第1のデータ記憶装置中で前記データ変更手段が行う変更を識別するための変更記録手段と、
- D. 前記変更記録手段中に記録された変更に従って、第1の動作モードの間に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モードを確立するための第1の動作手段と、
- E. 第2の動作モードが確立されるときに前記変更記録手段中に記録される変更に従って、前記第2のデータ記憶装置から前記受信手段にデータをコピーするための前記第2の動作モードを確立するための第2の動作手段と、を有することを特徴とするデータ記憶設備。

【請求項2】 前記データ記憶設備内の転送は、データブロックとして起こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の情報のための記憶装置を含み、前記変更記録手段は、前記別の情報を受信するための第1のテーブルを含み、前記第2の動作手段は、所定の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送することを特徴とする請求項1に記載のデータ記憶設備。

【請求項3】 前記データ記憶設備内の転送は、データブロックとして起こり、前記第1のデータ記憶装置の部分は、それぞれのデータブロックについての情報を含み、前記第2の動作手段は、前記第1のデータ記憶装置の部分の内容を前記変更記録手段に転送することを特徴とする請求項1に記載のデータ記憶設備。

【請求項4】 前記データ記憶設備は、前記第1及び第2の動作手段を相互に排他的に使用可能にする前記変更手段から受信する所定のデータに応答する制御手段を更に含むことを特徴とする請求項3に記載のデータ記憶設備。

【請求項5】 前記データ記憶設備は、カスケードセッションの間中に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可能にする、そのようなカスケードセッションを実行するための前記変更手段からの、反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に含むことを特徴とする請求項3に記載のデータ記憶設備。

【請求項6】 データを転送するためのコピー手段を更に含み、前記変更記録手段は、第1、第2及び第3のテーブルを含み、前記第1の動作手段は、所定の位置から前記第1のテーブルヘデータを転送し、かつ、カスケードセッションにおける最初の反復の間中に、前記第2のテーブル中の全てのデータを第1の状態にセットし、それによって前記コピー手段が前記第2のデータ記憶装置中の全ての前記データを前記受信手段に転送することを特徴とする請求項5に記載のデータ記憶設備。

【請求項7】 次の動作における前記第1の動作手段は、前記第1及び第2のテーブル中の対応するデータを変更し、次の動作における前記第2の動作手段は、前記第1のテーブルから前記第3のテーブルに対応するデータを転送することを特徴とする請求項6に記載のデータ記憶設備。

【請求項8】 ホスト及びプロダクション記憶設備を有するプロダクションサイトを含み、かつ、リモート記憶設備、前記プロダクションサイトを相互接続するための記憶設備、及び前記ホストが複数のカスケードコマンドを発行することができるような前記リモートサイトを含むデータネットワークにおける、中間記憶設備であって、

- A. トラック毎にデータを前記プロダクション設備から受信するための、前記プロダクション設備との接続のための第1のデータ記憶装置と、
- B. 第2のデータ記憶装置と、
- C. 前記プロダクション設備が変更する前記第1のデータ記憶装置中のそれぞれのトラックを識別するための複数のトラックステータステーブルと、
- D. 第1のカスケードコマンドに応答して、前記トラックステータステーブル中に記録された変更に従って、第1の動作モードの間中に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モ

ードを確立するための第1の動作手段と、

E. 第2のカスケードコマンドに応答して、第2の動作モードが確立されるときに前記トラックステータステーブル中に記録される変更に従って、前記第2のデータ記憶装置から前記リモート記憶設備にデータをコピーするための、前記第2の動作モードを確立するための第2の動作手段と、を有することを特徴とする中間記憶設備。

【請求項9】 前記第1のデータ記憶装置は、それぞれのトラックについての別の情報のための記憶装置を含み、前記トラックステータステーブルは、前記別の情報を受信するための第1のテーブルを含み、前記第2の動作手段は、所定の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送することを特徴とする請求項8に記載の中間記憶設備。

【請求項10】 前記第1のデータ記憶装置の部分は、それぞれのデータトラックについての情報を含み、前記第2の動作手段は、前記第1のデータ記憶装置の部分の内容を前記トラックステータステーブルに転送することを特徴とする請求項8に記載のデータ記憶設備。

【請求項11】 前記中間記憶設備は、前記第1及び第2の動作手段を相互に排他的に使用可能にする前記プロダクション設備から受信する所定のデータに応答する制御手段を更に含むことを特徴とする請求項10に記載のデータ記憶設備。

【請求項12】 前記中間記憶設備は、カスケードセッションの間中に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可能にする、そのようなカスケードセッションを実行するための前記プロダクション設備からの、反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に含むことを特徴とする請求項10に記載のデータ記憶設備。

【請求項13】 データを転送するためのコピー手段を更に含み、前記トラックステータステーブルは、第1、第2及び第3のテーブルを含み、前記第1の動作手段は、前記第1のデータ記憶装置の部分から前記第1のテーブルへデータを転送し、かつカスケードセッションにおける最初の反復の間中に、前記第2のテーブル中の全てのデータを第1の状態にセットし、それによって前記コピー手

段が前記第2のデータ記憶装置中の全ての前記データを前記受信手段に転送することを特徴とする請求項12に記載のデータ記憶設備。

【請求項14】 次の動作における前記第1の動作手段は、前記第1及び第2のテーブル中の対応するデータを変更し、次の動作における前記第2の動作手段は、前記第1のテーブルから前記第3のテーブルに対応するデータを転送することを特徴とする請求項13に記載のデータ記憶設備。

【請求項15】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するための手段を更に有することを特徴とする請求項10に記載のデータ記憶設備。

【請求項16】 前記中間記憶設備は、前記プロダクション設備から地理的に遠隔であり、前記プロダクション設備と前記中間記憶設備の間の通信リンクを確立するための手段を更に有することを特徴とする請求項10に記載のデータ記憶設備。

【請求項17】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するための手段を更に有することを特徴とする請求項16に記載のデータ記憶設備。

【請求項18】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するための手段を更に有し、前記通信リンクのそれぞれは、異なる特性のバンド幅を有することを特徴とする請求項16に記載のデータ記憶設備。

【請求項19】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するための手段を更に有し、前記プロダクション設備への前記通信リンクは、前記リモートサイトへの前記通信リンクより広い特性のバンド幅を有することを特徴とする請求項16に記載のデータ記憶設備。

【請求項20】 データ変更設備とデータ受信設備の間でデータを転送するための方法であって、

A. 変更されたデータを受信するための第1のデータ記憶装置を確立するステップと、

- B. 第2のデータ記憶装置を確立するステップと、
- C. 受信された変更されたデータによって作り出される前記第1のデータ記憶装置中の変更を識別するステップと、
- D. 識別された前記変更に従って、第1の動作モードの間に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モードを確立するステップと、
- E. 第2の動作モードが確立されるときに識別される変更に従って、前記第2のデータ記憶装置から前記受信設備にデータをコピーするための第2の動作モードを確立するステップと、を有することを特徴とする方法。

【請求項21】 前記データ記憶設備内の転送は、データブロックとして起こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の情報のための記憶装置を含み、前記変更を識別するステップは、前記別の情報を受信するための第1のテーブルを確立するステップ、及び前記第2の動作モードが確立しているときに所定の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送することを特徴とする請求項20に記載の方法。

【請求項22】 前記データ記憶設備内の転送は、データブロックとして起こり、前記第1のデータ記憶装置の部分は、それぞれのデータブロックについての情報を含み、前記第1のデータ記憶装置の部分の内容は、最初の識別情報を提供するために転送されることを特徴とする請求項20に記載の方法。

【請求項23】 前記第1及び第2の動作モードは、前記変更設備からの所定のデータに応答して相互に排他的に動作することを特徴とする請求項22に記載の方法。

【請求項24】 前記変更設備からの第1及び第2のコマンドの反復する生成は、カスケードセッションの間中に前記第1及び第2の動作モードが相互に排他的に反復して使用可能にされる、そのようなカスケードセッションを実行することを特徴とする請求項22に記載の方法。

【請求項25】 データは、コピーによって転送され、変更の前記識別は、第1、第2及び第3のテーブルを含み、第1の動作モードの開始は、所定の位置から前記第1のテーブルヘデータを転送し、かつカスケードセッションにおける

最初の反復の間中に、前記第2のテーブル中の全てのデータを第1の状態にセットし、それによって前記コピーが前記第2のデータ記憶装置中の全ての前記データを前記受信設備に転送することを更に特徴とする請求項24に記載の方法。

【請求項26】 次の動作における前記第1の動作モードは、前記第1及び第2のテーブル中の対応するデータを変更し、次の動作における前記第2の動作モードは、前記第1のテーブルから前記第3のテーブルに対応するデータを転送することを特徴とする請求項25に記載の方法。

【請求項27】 ホスト及びプロダクション記憶設備を有するプロダクションサイトと、リモート記憶設備を有するリモートサイトの間でデータを転送するための方法であって、前記ホストは、複数のカスケードコマンドを発行することができ、

A. トラック毎にデータを前記プロダクション設備から受信するための第1のデータ記憶装置を確立するステップと、

B. 第2のデータ記憶装置を確立するステップと、

C. 前記プロダクション設備が変更する前記第1のデータ記憶装置中のそれぞれのトラックを識別するための複数のトラックステータステーブルを定義するステップと、

D. 第1のカスケードコマンドに応答して、前記トラックステータステーブル中に記録された変更に従って、第1の動作モードの間中に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モードを確立するステップと、

E. 第2のカスケードコマンドに応答して、第2の動作モードが確立されるときに前記トラックステータステーブル中に記録される変更に従って、前記第2のデータ記憶装置から前記リモート記憶設備にデータをコピーするための、前記第2の動作モードを確立するステップと、を有することを特徴とする方法。

【請求項28】 前記第1のデータ記憶装置は、それぞれのトラックについての別の情報のための記憶装置を含み、前記変更トラックステータステーブルは、前記別の情報を受信するための第1のテーブルを含み、前記第2の動作手段は、所定の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送

することを特徴とする請求項27に記載の方法。

【請求項29】 前記第1のデータ記憶装置の部分は、それぞれのトラックについての情報を含み、前記第2の動作手段は、前記第1のデータ記憶装置の部分の内容を前記トラックステータステーブルに転送することを特徴とする請求項27に記載の方法。

【請求項30】 前記中間データ記憶設備は、前記第1及び第2の動作手段を相互に排他的に使用可能にする前記プロダクション設備から受信する所定のデータに応答する制御手段を更に含むことを特徴とする請求項29に記載の方法。

【請求項31】 前記中間データ記憶設備は、カスケードセッションの間中に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可能にする、そのようなカスケードセッションを実行するための前記プロダクション設備からの、反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に含むことを特徴とする請求項29に記載の方法。

【請求項32】 データ転送を実行するためにデータをコピーし、第1、第2及び第3のテーブルを有する前記トラックステータステーブルを確立し、前記第1のデータ記憶装置の部分から前記第1のテーブルヘデータを転送し、かつカスケードセッションにおける最初の反復の間中に、第1の動作モードの第1の状態に前記第2のテーブル中のすべてのデータをセットし、それによって前記コピー手段が前記第2のデータ記憶装置中の全ての前記データを前記受信手段に転送するステップを更に含むことを特徴とする請求項31に記載の方法。

【請求項33】 次の反復における前記第1の動作モードは、前記第1及び第2のテーブル中の対応するデータを変更し、次の反復における前記第2の動作モードの間に、前記第1のテーブルから前記第3のテーブルに対応するデータが転送されることを特徴とする請求項32に記載の方法。

【請求項34】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するステップを更に有することを特徴とする請求項29に記載の方法。

【請求項35】 前記中間記憶設備は、前記プロダクション設備から地理的に遠隔であり、前記プロダクション設備と前記中間記憶設備の間の通信リンクを

確立するステップを更に有することを特徴とする請求項 29 に記載の方法。

【請求項 36】 前記中間記憶設備は、前記リモートサイトから地理的に遠隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立するステップを更に有することを特徴とする請求項 35 に記載の方法。

【請求項 37】 前記通信リンクを確立する前記ステップは、異なる特性のバンド幅を有する通信リンクを確立することを特徴とする請求項 35 に記載の方法。

【請求項 38】 前記通信リンクを確立する前記ステップは、前記リモートサイトへの前記通信リンクより広い特性のバンド幅を有する前記プロダクション設備への通信リンクを確立することを特徴とする請求項 35 に記載の方法。

【発明の詳細な説明】

【0001】

(関連出願とのクロス・リファレンス)

出願がこの発明と同じ譲受人に譲渡された、1996年5月31日に出願されたりリモートデータ設備の独立動作の方法及び装置(Method and Apparatus for Independent Operation of a Remote Data Facility)のための米国特許第08/656,035号。

【0002】

出願がこの発明と同じ譲受人に譲渡された、1997年4月25日に出願された共通データセットへの独立かつ同時のアクセスのための方法及び装置(Method and Apparatus for Independent and Simultaneous Access to a Common Data Set)のための米国特許第08/656,035号。

【0003】

(技術分野)

この発明は、広くデータ処理ネットワーク中の冗長データ記憶デバイスに関し、より詳しくは、複数の冗長データ記憶ユニット(redundant data storage unit)を通じてデータをカスケードにする(cascade)ことを可能にする方法及び装置に関する。

【0004】

(背景技術)

データの冗長性によるデータの健全性の維持は、非常に重要な問題となっている。データの冗長性は、いくつかの形態及びバリエーションを有する。単一のサイトでは、ミラリング(mirroring)又はRAIDの冗長性が、ディスク又は他の記憶装置の故障から保護する。1996年5月31日に出願された米国特許第08/656,035号などに記載された他の形態では、冗長性は、遠隔の位置にあるローカルなシステムを複製することによって達成される。遠隔の場所を使用することによって、1つの場所での自然災害などによってデータが失われることが防止される。

【0005】

先の米国特許第08/656,035号によると、すべてのデータ処理活動は、データ記憶設備中に記憶されたデータを処理するためのホストシステムを含む「ローカル」の又は「プロダクション(production)」のサイトで起こる。地理的に遠隔な又はバックアップのサイトは、復元(restoration)又は回復(recovery)データセットを維持するための「冗長な」設備として、データ記憶設備を含む。このシステムでは、プロダクションサイトでホストがプロダクションサイトデータ記憶設備にデータを書き込む度に、プロダクションデータ記憶設備は、リモート記憶設備にデータを自動的に書き込む。これらのアプリケーションの、ほとんどではないとしても多くでは、リモートサイトのデータ記憶設備にデータを書き込むことは、バックアップ手順がプロダクション設備の運転に影響を与えないように、広いバンド幅の通信リンクを通した転送を必要とする。T3及びESCONラインは、それらの費用にもかかわらず、通信リンクとして、典型的に好まれている。

【0006】

米国特許第08/842,953号は、プロダクションサイトで典型的に使用されるBCVデバイスを含む他の概念を開示する。このアプローチでは、データセット、例えば、プロダクションサイトのデータ記憶設備上の「論理ボリューム(logical volume)」、すなわち「プロダクションボリューム(production volume)」は、通常の方法で動作する。この同じサイトでの他の論理ボリュームは、「BCVボリューム(BCV volume)」として示される。ESTABLISHコマンドは、BCVボリュームをプロダクションボリュームに接続し、BCVボリュームは、プロダクションボリュームと同期する。SPLITコマンドは、その後、BCVボリュームをプロダクションボリュームから分離することができ、BCVボリューム上に記録されたデータを他のアプリケーションで利用できるようにすることができる。

【0007】

この他のアプリケーションは、BCVボリューム上に記憶されたデータを変更することができる。結果として、BCVボリュームがプロダクションボリュームと再接続されるときはいつでも、BCVボリュームとプロダクションボリュームの両方の変化を表わすデータをBCVボリュームに転送する必要がある。他の選択肢は、

別のESTABLISHコマンドを出すことによって、BCVボリュームをプロダクションボリュームに再接続することである。しかし、このコマンドは、BCVボリューム上のすべてのデータを置換するであろう。変化が起こったかどうかにかかわらず、プロダクションボリュームからのすべてのデータが、BCVボリュームに転送されなければならない。

【0008】

過去では、プロダクションサイトでのデータ記憶設備の位置を示すことは、許容されてきた。特定の重要な用途では、第1の2つのサイトから遠隔の第3のサイトで更に他の復元コピーを記憶させることが今でも望まれ、またいくつかの状況では、それは義務的である。これによって、すべてプロダクションサイトの動作に透過的に(transparently)、プロダクションサイトから両方のリモートサイトにデータをコピーするためのアプローチが必要となる。プロダクションサイトから単一のリモートサイトにデータをコピーするための前述のアプローチを単に反復することは、過度の通信コストを含む。プロダクションサイトと第2のリモートサイトの間、あるいは第1と第2のリモートサイトの間のいずれかで、第2の高速通信リンクが必要となるであろう。必要なものは、よりコストの低い通信リンクを通じた、種々の遠隔に位置するサイトで存在するデータに重大な相違が全くない、すべてのデータ記憶設備が同期するか又はほぼ同期するような、冗長なデータのコピーを確立する方法である。

【0009】

(発明の開示)

従って、本発明の目的は、よりコストの低い通信リンクを通じて通信することができる冗長なデータ記憶設備を確立するための方法及び装置を提供することである。

【0010】

本発明の他の目的は、地理的に遠隔なサイトで、冗長なディスク記憶を実現するための方法及び装置を、経済的に提供することである。

【0011】

本発明の更に他の目的は、経済的な通信リンクを使用して、2つ以上の地理的

に遠隔なサイトに、冗長なデータ記憶を提供するための方法及び装置を提供することである。

【0012】

本発明のまた更に他の目的は、経済的な低いバンド幅の通信リンクを使用して、2つ以上の地理的に遠隔なサイトに、冗長なデータ記憶を提供するための方法及び装置を提供することである。

【0013】

本発明のまた更に他の目的は、プロダクションサイトで処理されるデータが、プロダクションサイトでの動作に透過的に達成される冗長性で、2つ以上の冗長な記憶サイトに記憶されるデータ処理ネットワークを提供することである。

【0014】

本発明によると、プロダクション設備のような、データを変更するデータ処理設備から、リモートデータ記憶サイトは、データ記憶設備から遠隔の他のサイトにデータを転送する。リモートサイトのデータ記憶設備は、第1及び第2のデータ記憶装置(data store)を含む。第1のデータ記憶装置は、プロダクションサイトからデータを受信する。データ変更レコーダ(data change recorder)は、データ処理設備が第1のデータ記憶装置中で行った変更を識別する。第1の動作コントロール(operating control)は、その動作モードの間に前記変更レコーダ手段中で記録された変更に従って第1のデータ記憶装置からデータを受信する、そのような第1の動作モードを確立する。この動作が完了した後、第2の動作コントロールは、前記データ変更レコーダ中に記録された変更に従って、第2のデータ記憶装置から他のリモートサイトにデータをコピーするための第2の動作モードを確立する。

【0015】

本発明の他の態様によると、データは、プロダクションサイトとリモートサイトの間で転送される。プロダクションサイトは、ホスト及びプロダクション記憶設備を含み、リモートサイト、リモート記憶設備は、トラック毎に(track-by-track basis)プロダクション設備からデータを受け取ることを特徴とする第1のデータ記憶装置、及び第2のデータ記憶装置を含む。ホストは、プロダクション設

備が変更する第1のデータ記憶装置中のそれぞれのトラックを識別するための、複数のトラックステータステーブルの画定を可能にする、複数のカスケードコマンドを発行することができる。更に、ホストは、第1及び第2の動作モードを確立することができる。第1の動作モードでは、第2のデータ記憶装置は、トラックステータステーブル中に記録された変更に従って、第1のデータ記憶装置からデータを受信する。第2の動作モードの間では、第2のデータ記憶装置からのデータは、第2のモードが確立されるときに、トラックステータステーブル中に記録される変更に従って、リモート記憶設備にコピーされる。

【0016】

(図面の簡単な説明)

添付の請求項は、本発明の主題を、特に指し示し、明確にクレームする。本発明の種々の目的、利点及び新しい特徴は、同様の参照番号は同様の部品を示す、そのような添付の図面と共に、以下の詳細な説明を読むことから、より完全に明らかになるであろう。

【0017】

(本発明を実施するためのベストモード)

(一般的動作)

図1は、ローカル又はプロダクションサイト21、第1のリモートサイト22及び第2のリモートサイト23に装置が設置されたデータ処理ネットワーク20を示す。第1及び第2のリモートサイト22及び23は、典型的には、ローカルのプロダクションサイト21から、及びお互いから、地理的に離れているであろう。しかし、明らかになっていくように、ローカルのプロダクションサイト21及び第1のリモートサイト22の装置は、一緒に置くことができるであろう。

【0018】

冗長性の第1のレベルは、ローカルのプロダクションサイト21と第1のリモートサイト22の間の相互作用(interaction)を通じてデータ処理ネットワーク20中で達成される。知られているように、1つ以上の中央プロセッサ及びメインメモリを含むホスト24は、種々のプログラムで動作する。周期的に、ホスト24は、ホストアダプタ25を通じてディスク記憶デバイスへの転送を実行する

であろう。このディスク記憶デバイスは、関連する情報を記憶するための別個のセクションに編成された多くの物理ディスクドライブを有することができる。これらは、ファイル又は他のデータブロックを含む。本発明の譲受人によって製造される装置の状況では、典型的な記憶セクションは、多くの隣接したディスクトラックを含む論理ボリュームであり、転送はトラック毎に実行される。図1では、R1論理ボリューム26は、通常そのようなディスク記憶設備中に含まれる多くの論理ボリュームを表わす。明らかになっていくように、ホスト24は、それがそのようなプロダクション記憶設備からホストアダプタ25を通じて必要とするどのような情報も検索する。そのようなシステムは、技術上周知である。米国特許第08/656,035号は、1つのそのようなシステムを開示する。

【0019】

示されていないが、ローカルのプロダクションサイト21は、R1論理ボリューム26のための冗長性を提供することができる。例えば、R1ボリュームを、実際には、ミラリングしたり、又は種々のRAID構成によってホスト24による割込み動作から特定の物理ディスクドライブの機能不全を防止する、そのような種々のRAID構成のいずれか1つで、論理ボリュームのアレイによって構成したりすることができる。明らかになっていくように、ローカルのプロダクションサイト21で実施されるどのような冗長性の構成にも関わらず、本発明を適用することができる。

【0020】

技術上知られ、本発明と同じ譲受人に譲渡されたヤナイ他への米国特許第5,544,347号に記載されているように、第1のリモートサイト22でR1論理ボリューム26中のデータをミラリングすることによって実施することができる他の冗長性の構成が開示される。この構成においてより具体的には、ローカルのプロダクションサイト21は、ESCON又はT3通信ラインのような高速通信リンクを通じて、第1のリモートサイト22中の対応するリモートアダプタ31に接続するリモートアダプタ(remote adapter, RA)27を含む。この構成では、ホスト24がR1論理ボリューム26にデータを書き込むごとに、R2論理ボリューム32への転送のためにそのデータを高速通信リンクを通じて第1のリモート記憶サイ

ト22中のリモートアダプタ31に転送することによって、リモートアダプタ27は応答する。このように、それぞれのWRITE(書込み)動作は、R1論理ボリューム26に関して遠隔のミラーとしての役割をするR2論理ボリューム32に反映すなわち伝達される。もし、自然災害がローカルのプロダクションサイト21を襲ったなら、データは、リモートサイトのR2論理ボリュームで簡単に利用できる。更に、もしホスト33が第1のリモートサイト22中に位置するなら、すべての動作は、検出されないデータの損失が全くないまま、連続した動作のために第1のリモートサイトにシフトすることができる。

【0021】

前述のように、米国特許出願第842,953号は、データ記憶設備に関するBCV論理ボリュームを含むデータ処理ネットワークを開示する。本発明によると、BCV/R1論理ボリューム34は、第1のリモートサイト22中に含まれる。それは、第1のリモートサイト22内の、好適にはR2論理ボリューム32を含む物理ディスクドライブと異なる物理ディスクドライブ上に、任意の専用の論理ボリュームを含むことができる。

【0022】

この発明によると、また後でより詳細に説明するように、BCV/R1論理ボリューム34を、R2論理ボリューム32又はリモートアダプタ35のいずれかに接続することができる。第1の動作モードでは、BCV論理ボリューム34は、R2論理ボリューム32と同期する。BCV/R1論理ボリューム34がリモートアダプタ35に接続している第2の動作モードでは、データは、R2論理ボリューム40又は他のデータレシーバへの転送のため、他の通信リンク36を通じて、第2のリモートサイト23中のリモートアダプタ37に転送されるであろう。本発明の1つの実施形態では、第2のリモートサイト23は、ホスト41を含み、また第1のリモートサイト22にホスト33が存在する必要性をなくす。このようにして、第2のリモートサイト23は、もし自然災害がローカルのプロダクションサイト21で起きると、復元サイト又はデータで動作するための第2のサイトになる。

【0023】

第2のリモートサイトは、オプションのBCV/R1論理ボリューム42を含むよう

に示される。後でより完全に説明するように、第2のリモートサイト23にこの論理ボリュームを含ませることによって、第3のリモートサイトへの本発明に従って実施される機能の複製(replication)ができる。

【0024】

一般的な条件では、本発明に従って構成されるリモートサイトは、データ記憶設備を構成する第1のリモートサイト22の基本的な構造を有するであろう。それは、データを変更することができるローカルのプロダクションサイト21への接続のために、R2論理ボリューム32の形態の第1のデータ記憶装置を含む。BCV/R1論理ボリューム34は、第2のデータ記憶装置を構成する。第1のコマンドに応答して、第1のリモートサイト22中のデータ記憶設備は、その動作モードの間にR2論理ボリューム32がリモートアダプタ31からデータを受信し、それによってR1論理ボリューム26中のデータになされた変更に応答する、そのような第1の動作モードで動作する。典型的には、これは、同期的に実施され、そのため高速通信リンク30が必要である。この動作モードでは、BCV/R1論理ボリューム34は、そのBCV又は第1の動作モードで動作していると考えられる。

【0025】

本発明によると、BCV/R1論理ボリューム34は、一旦、第1の動作モードで同期が達成されると、第2の又はR1の動作モードにシフトすることができる。第2の動作モードでは、第1のリモートサイトは、BCV/R1論理ボリューム34からリモートアダプタ35、通信リンク36及びリモートアダプタ37を通じて、R2論理ボリューム40にデータを転送する。第1から第2の動作モードへのシフトのタイミングは、システムのエオペレータによって決定されるであろう。しかし、第2の動作モードから第1の動作モードへのシフトは、一般的には、R2論理ボリューム40へのデータ転送が完了した後に実施される。典型的には、第2の動作モードへのシフトの間の間隔は、R2論理ボリューム32中の活動によって、分、時間又は更には日の期間になるであろう。

【0026】

更に本発明によると、BCV/R1論理ボリューム34が、その第2の動作モードにシフトするたびに、第1の動作モードの間に変更されたデータトラックのみが

、R2論理ボリューム40に転送される。もしローカルのプロダクションサイト21が、第2の動作モードへの連続したシフトの間にR2論理ボリューム32の単一のトラックへの反復した変更をするなら、BCV/R1論理ボリューム34からR2論理ボリューム40への1つの転送のみが起こるであろう。そのような減少によって、通信リンク36へのバンド幅の要件を減少させることができる。例えば、通信リンク36を、低いバンド幅の電話線又はインターネットを通じて転送することができるようなレベルに減少させることができるであろう。

【0027】

このように、本発明に従って、R1論理ボリューム26になされたデータの変更は、R2論理ボリューム32中で、そしてBCV/R1論理ボリューム34を通じてR2論理ボリューム40へと複製される。順番にデータを転送するこの処理は、カスケードとして特徴付けられ、ローカル又はプロダクションサイト21でホスト24によって処理されるカスケードコマンドセットによって制御される。

【0028】

(カスケードコマンド処理)

より具体的には、システムのオペレータ、又はローカルのプロダクションサイト21で処理されているアプリケーションプログラムは、第1のリモートサイト22で実行されているカスケード動作を実行する。知られているように、図1のホスト24は、カスケードコマンドセット中のコマンドを含む種々のコマンドを作り出すであろう。これらのコマンドは、ホストアダプタ25を含む任意の数のデバイス又はアドレスに命令することができる。

【0029】

ここで図1及び2を参照すると、ホストアダプタ25がコマンドを受信するとき、ステップ44は、そのコマンドをテストしてデコードするために普通の処理を使用する。そのテストは、正しいコマンドが受信されたかどうかを確定するために、文法、コンテキスト(context)及び他の関連するパラメータを分析する。ステップ45は、そのコマンドが、カスケードコマンドセット中のコマンドの1つであるかどうか、すなわち、カスケードコマンドであるかどうかを確定する。もし、それがカスケードコマンドでなければ、ホストアダプタ25は、対応する

機能を実施するために手順46を使用する。有効なカスケードコマンドを受信すると、ホストアダプタ25は、そのコマンドをリモートアダプタ27に転送するために、ステップ47を使用する。

【0030】

リモートアダプタ27がステップ50でコマンドを受信したとき、ステップ51は、それ以上分析することなく、そのコマンドを通信リンク30を通じてリモートアダプタ31に、単に転送する。リモートアダプタ31は、同様の処理を実行する。すなわち、リモートアダプタ31は、ステップ52でコマンドを受信し、及びそのコマンドをデバイスコントローラ43に転送するためにステップ53を使用する。

【0031】

デバイスコントローラ34が、ステップ54でリモートアダプタ31からコマンドを受信した後、デバイスコントローラ34は、ステップ55でコマンドをデコードし、及びステップ56でそのコマンドを処理する。

【0032】

カスケードコマンド又は他のコマンドによって定義された動作が完了したとき、デバイスコントローラ43は、リモートアダプタ31への転送のために、ステップ60で受取り(acknowledgement)を生成し、ステップ61、62及び63は、ホスト24への転送のために、その受取りを通信リンクを逆に通してリモートアダプタ27及びホストアダプタ25に転送するための処理を表わす。

【0033】

このように、この処理によって、ホスト24中で動作しているプログラムが、今度はBCV/R1論理ボリューム34の動作を制御する第1のリモートサイト中のデバイスコントローラ43へのコマンドを発行することができる。ステップ60から63は、次に、ホストがすべてのそのような動作の完了に気づくように、必要なフィードバックを提供する。結果として、本発明を実施するための第1のリモートサイトで、図1のホスト33のようなホストに対する要件がなくなる。

【0034】

(第1のリモートサイト22)

図1をもう一度参照すると、第1のリモートサイト22は、本発明を実施するための、多くのステータスレジスタ及びテーブル及びプログラムモジュールを含む。コントロールモジュール70は、ESTABLISHコマンドを処理するためのESTモジュール71、DIFFERENTIAL SPLITコマンドを処理するためのDSモジュール72、及びREESTABLISHコマンドを処理するためのREESTモジュール73を含む、カスケードコマンドセット中のコマンドを含む種々のコマンドを処理するためのモジュールを含む。ESTABLISH及びREESTABLISHコマンドは、カスケード動作から独立して、機能も実行する。

【0035】

論理ボリュームのそれぞれは、デバイスのための、個々の物理シリンダー、及びそれらのシリンダー内のトラックのステータスについての情報を維持するための専用の部分を含む。R2論理ボリューム32の1つのそのような部分は、それぞれのトラックについてのステータスワードを含む専用の部分74によって構成される。それぞれのステータスワードは、トラックについての情報を含み、種々のステータスの機能について利用可能な多くのビット位置(bit position)のために使用される。それらのビット位置の1つは、カスケード処理での使用のために割り当てられるであろうし、またそのビット位置は、nがR2論理ボリューム32中のトラックの数と等しい場合の $1 \times n$ のテーブルである、プロテクションビットテーブル(Protection Bit Table)75として複製されるであろう。

【0036】

第1のリモートサイト22は、R2トラックステータステーブル76及びBCV/R1トラックステータステーブル77も含む。トラックステータステーブル76及び77のそれぞれは、複数の行、具体的には、4つのミラリングデバイス又はボリュームに割り当てられた4つの行を含む。BCV/R1論理ボリューム34がR2論理ボリューム32に接続しているときの通常動作では、R2トラックステータステーブル76のM1行中のビット位置は、R2論理ボリューム32中のトラックに対応し；M2行中のビット位置は、R1論理ボリューム26中のトラックに対応し；及びM3行中のビット位置は、BCV動作モードで働いているBCV/R1論理ボリューム34中のトラックに対応する。同様に、BCV/R1トラックステータステーブル77は、BCV/

R1論理ボリューム34のためのトラックステータスを識別するためのM1行中のビット位置及びR2論理ボリューム40のためのM2行中のビット位置を含む複数のミラー行を含む。

【0037】

PBビットレジスタ80は、カスケード動作の間に使用されることになるプロテクションビットテーブル75中の選択されたビット位置を識別する。セッションフラグ81は、システムがカスケードコマンドに応答して動作しているかどうかを指示する。コピープログラム82は、後でより詳細に説明するように、第1のリモートサイト22設備中の異なる論理ボリュームの間でのデータのコピーを行う。

【0038】

(動作-第1の動作モード)

第1のリモートサイト22は、リモートアダプタ31で受信された変更に従って、その動作モードの間にコピープログラム82がR2論理ボリューム32の形態の第1のデータ記憶設備からBCV/R1論理ボリューム34の形態の第2のデータ記憶設備にデータをコピーする、そのような第1の動作モードで動作するように、最初にコンフィギュレーション設定が行われるであろう。このコンフィギュレーションは、BCV/R1論理ボリューム34上のデータがR2論理ボリューム32上に記憶されたデータと等しくなるような時間まで、同じ状態を保つ。その後、第1の動作モードは、その同期を維持する。

【0039】

(動作-第2の動作モード)

そのような同期が達成された後はいつでも、ホスト24がDIFFERENTIAL SPLITカスケードコマンドを発行するときに、第2の動作モードを確立することができる。他には、第1のリモートサイトのオプションのホスト33は、そのようなコマンドを生成することも可能であろう。コマンドを発行することは、カスケードセッションを開始する。

【0040】

応答として、デバイスコントローラ43は、BCV/R1論理ボリューム34をR2論

理ボリューム32から分離し、図1のコピープログラムがBCV/R1論理ボリューム34からリモートアダプタ35を通じて、R2論理ボリューム40を含む図1の第2のリモートサイト23によって表わされるようなデータ受信デバイスにデータを転送することを可能にするであろう。それらの2つのボリュームが同期した後はいつでも、BCV/R1論理ボリューム34をリモートアダプタ35から分離し、それをR2論理ボリューム32に再接続することによって、REESTABLISHコマンドはシステムを第1の動作モードに戻し、その後、コピープログラム82は、BCV/R1論理ボリューム34をR2論理ボリューム32と再同期させるであろう。セッションは、後続のDIFFERENTIAL SPLIT及びREESTABLISHコマンドを発行することによって、CLOSE SESSIONコマンドが生成されるまで継続する。

【0041】

図3は、より詳細にDIFFERENTIAL SPLITコマンドの受信に応答した動作を示す。ホストアダプタ25が、ステップ90で異なるコマンドが受信されたことを認識するとき、種々のあり得るエラーステータスを求めて、知られているように及び図2のステップ44に関して説明されるように、コマンドをテストする。もし何かエラーが存在するなら、その手順は、技術上知られた報告手順を使用して中止される。もしエラーが存在しないなら、制御部(control)70が応答できるように、リモートアダプタ27、通信リング30及びリモートアダプタ31を通じた転送のために、DIFFERENTIAL SPLITコマンドを転送するステップ92に制御が移る。

【0042】

DIFFERENTIAL SPLIT(DS)モジュール72がステップ93で応答するとき、それは、ロック(lock)が存在する間は、他のどのアプリケーションも選択された論理ボリュームに影響を与えることを防ぐために、普通のロック動作(locking operation)を実行する。ステップ94は、セッションフラグ81をモニタし、カスケードセッションが進行中であるかどうかを確認する。もしこれが、カスケード動作の最初のディファレンシャルスプリット(differential split)動作であれば、セッションフラグ81は、セッションは進行中でないことを指示する。制御は、所定の位置に記憶されたデータに基づいてプロテクションビットテーブル(prote

ction bits table) 75を確立するステップ95に進行する。例えば、ステップ95は、R2論理ボリューム32中の専用の部分74から、プロテクションビットデータファイル(protection bits data file)を検索することができる。他には、ステップ95は、メモリ中に維持されているそのようなファイルのコピーから、対応する情報を検索することができる。どのような形態でも、それぞれのエントリは、トラックに対応しており、それぞれのエントリは、任意の目的に割り当てることができる複数のビット部分を含む。ステップ95は、開始されているセッションのための1つのビット位置を確立するため、異なるビット位置のどのような有効な使用も識別するテーブル(図示せず)から、使用されていないビット部分を選択する。そのようなビット位置を選択するための処理は、技術上周知である。次に、DSモジュール72は、その選択されたビット位置をPB BITレジスタ80中に記録する。

【0043】

ステップ96は、普通のモジュールから基本的なステップを使用して普通のスプリット動作を実行するための手順を表わす。この動作の詳細は、図4に示される。具体的には、ステップ101は、R2論理ボリューム32のためのミラーとしてのBCV/R1論理ボリューム34のミラリング動作を終了させる。この実行によって、コピープログラム82が、R2論理ボリューム32からBCV/R1論理ボリューム34に転送されることが防止される。しかし、R2論理ボリューム32へのそれぞれの転送は、R2トラックステータステーブル76のM3行中に記録され、それによって、BCV/R1論理ボリューム34中に反映されなかったR2論理ボリューム26中の変化が指示されるであろう。

【0044】

ステップ102は、前述の米国特許出願第08/842,953号中に記載のように、WRITE PENDING動作を管理する。次のステップ103は、第1のリモートサイト22から第2のリモートサイトへのデータの転送がすべての識別情報を含むように、どの識別テーブル(identification table)も、R2論理ボリューム32からBCV/R1論理ボリューム34にコピーする。ステップ104は、次に、BCV/R1論理ボリューム34を、第2のリモートサイト23のR2論理ボリューム40

にデータをコピーするために割り当てる。これによって、コピープログラム82は、適当な宛先、すなわちR2論理ボリューム40にデータを転送することが可能になる。

【0045】

図4の動作が完了したとき、図3のステップ105は、BCV/R1トラックステータステーブル77のM2ビット位置中のすべてのビットをセットする。全てのビットをセットすることによって、すべてのトラック中のデータが、図1のR2論理ボリューム40に転送されることになる動作環境が確立される。ステップ105は、どのような後続のDIFFERENTIAL SPLITの間でも、後述の他のパスが続くことが明確になるように、セッションフラグ81もセットする。

【0046】

ステップ106は、ロックを解除する。ステップ107は、BCV/R1トラックステータステーブル77中のデータ、特にM3行中の情報に従って、BCV/R1論理ボリューム34からR2論理ボリューム40へのデータの転送を開始するコピープログラム82を開始する。それぞれのコピー動作が起きると、BCV/R1トラックステータステーブル77中の対応するM2ビットが消去される。全てのビットが消去されたとき、ステップ110は、完了ステータス(complete status)を通知(post)し、対応するメッセージは、ホストアダプタ25に逆に転送され、それによって、転送が完了したホスト24中で動作する対応するアプリケーションプログラムを指示する。

【0047】

(動作-第1の動作モードに戻る)

すべてのデータがR2論理ボリューム40に転送された後はいつでも、システムオペレータは、別のカスケードコマンドを構成するREESTABLISHコマンドを発行することができる。図5のステップ112は、ホストアダプタ25でのそのコマンドの受信を表わす。ステップ113は、もしエラーが検出されたなら処理を中止する可能性がある、そのコマンドのコンテキスト及び文法のテストを表わす。エラーが検出されなければ、ホストアダプタ25は、ステップ114を使用して、リモートアダプタ27及び通信リンク30を通じたりモートアダプタ31への

転送を制御するためのコマンドをデバイスコントローラ43に渡す。

【0048】

デバイスコントローラ43は、ステップ115を使用して、そのようなコマンドの受信を指示する。ステップ116は、BCV/R1論理ボリューム34を、R2トラックステータステーブル76中の次に利用可能なデバイスのミラーの指示を有するローカルのBCVミラーとして追加する。ステップ117で、デバイスコントローラ43は、DIFFERENTIAL SPLIT動作のための記憶設備として動作しているBCV/R1論理ボリューム34をNOT READY(NR)にセットする。これによって、COPYプログラムが更にデータを第2のリモートサイト23にコピーすることが防止される。どのWRITE PENDING動作も、ステップ118でINVALID状態にセットされる。前述のアプリケーションで説明したように、どのWRITE PENDING動作をセットすることも、記憶されているデータに影響を与えない。ステップ120は、次に、BCV/R1トラックステータステーブル77のためのM4ビット位置中のビットを有するBCV/R1論理ボリューム34（すなわち、M3ビット）に割り当てられたミラリングされた位置にあったR2トラックステータステーブル76中のどのビットもマージする。他のアプリケーションの間に、M4ビット位置は、BCV/R1デバイスのデータになされたどのような変更も、通常記録する。しかし、DIFFERENTIAL SPLITコマンドの間でシステムが第2のモードで動作しているとき、そのような変更はなされない。続いてステップ120は、第1のリモートサイトがDIFFERENTIAL SPLIT動作モードで動作していた間にR2論理ボリューム32になされたどのような変更も表わすR2トラックステータステーブル76のM3行中のビットを単に取得する。次のステップ121は、ロックを解除し、またステップ122は、ホストアダプタ25に逆に転送される完了ステータス(complete status)を、ステップ123でホスト24に通知する。ここで、コピープログラム82が、R2トラックステータスレジスタテーブル76中のM3ビット位置の内容に従って、R2論理ボリューム32からBCV/R1論理ボリューム34にデータを転送するために利用可能になる。それぞれの転送の間、コピープログラム82は、R2トラックステータスレジスタテーブル76のM3ビット位置中の対応するビットを消去する。しかし、プロテクションビットテーブル75中の対応するビット位置は、変化しないままである。

【0049】

(動作-R2論理ボリューム40への書き込み)

第1のリモートサイト22のデータ記憶設備が、BCV/R1論理ボリューム34からR2論理ボリューム40にデータをコピーする第2の動作モードで動作する間隔の間、ホスト24は、R1論理ボリューム26及びR2論理ボリューム32に書き込み動作を発行し続けることができる。リモートアダプタ31は、図6のステップ130でWRITEデータを受信する。その情報は、ステップ131でR2論理ボリューム32に直ちに書き込まれ、R2トラックステータステーブル76中の対応するエントリはステップ132で更新される。この場合、M3行中の対応するトラックビット位置は、WRITE動作がR2論理ボリューム32への転送を引き起こしたが、BCV/R1論理ボリューム34中の対応する変更を引き起こさなかったことを示すように更新されるであろう。次に、システムは、セッションフラグ81を見て、カスケード動作が進行中であったかどうかを確認するであろう。もしそうであったなら、ステップ133は、制御をステップ134に移し、それによって、プロテクションビットテーブル75及びトラックステータステーブル76中のM3行が同一の情報を有するように、対応するステータスをプロテクションビットテーブル75に書き込む。ステップ134は、もしカスケード動作が進行中でないなら、バイパスされる。ステップ135は、次に、その処理によって制御部70が書き込み動作を完了させる、そのような処理を表わす。通常動作の場合は、WRITE動作は、変更されたデータをBCV/R1論理ボリューム34に転送することによって完了する。もし、BCV/R1論理ボリューム34が接続されていなければ、WRITEリクエストは、WRITE PENDINGリクエストにされるであろう。

【0050】

(動作-連続ディファレンシャルスプリット動作)

システムのエベレータが定義することができ、分、時間又は日で測定することができる、いくつかの追加の時間間隔の後、ホスト21は、他のDIFFERENTIAL SPLITコマンドを発行する。デバイスコントローラ43がそのコマンドを受信するとき、その制御は、図3に示された手順に再び移される。しかし、この場合、図4のステップ101に関して説明したものと同様の方法で、ステップ94がス

ステップ141に分岐するようにセッションフラグ81はセットされ、R2論理ボリューム32へのミラーとしてのBCV/R1論理ボリューム34の動作を終了させる。これによって、R2論理ボリューム32からBCV論理ボリューム34へのデータの転送が目的の、コピープログラム82のどのような更なるアクションもできないようにされる。ステップ142は、ステップ102と同様に、全てのWRITE PENDING動作を管理する。

【0051】

次に、DSモジュール72は、ステップ143を使用して、プロテクションビットテーブル75を、BCV/R1トラックステータステーブル77中のM2行にコピーする。次に、DSモジュール72は、プロテクションビットテーブル75を消去する。これらの動作の後、BCV/R1トラックステータステーブル77中のM2ビット位置は、処理されている最後のDIFFERENTIAL SPLITリクエストに続いて変更されたR2論理ボリューム32中のそれらのトラックを指示する。プロテクションビットテーブル75中のプロテクションビットを消去することは、論理ボリューム32及び34のロックを外すと同時に、どのWRITE PENDING動作も、プロテクションビットテーブル75中の適当なビットをセットすることによって適切に記録されることを確実にする。

【0052】

DSモジュール72中の次の制御は、ステップ93でかけられたロックを解除するためにステップ143からステップ106にシフトし、そしてステップ107にシフトし、その結果、コピープログラム82が、BCV/R1論理ボリューム34からR2論理ボリューム40にデータを転送できるようになる。しかし、変更されたトラックの中のデータのみが送信されるであろう。すなわち、セッション中の最初の動作の後の全てのDIFFERENTIAL SPLIT動作は、変更されたトラック中のデータに対応する、増加する量の情報のみを送信するということは、このコマンドの性質である。更に、通信リンク36を通じて実際に転送されたデータのトラックの数は、連続したDIFFERENTIAL SPLIT動作の間に起こる書込み動作の数より、通常かなり小さいであろう。例えば、もしそのような間隔の間に、R2論理ボリューム32の同じトラック上の異なる領域への、10回の書込み動作があったとした

ら、BCV/R1デバイス34は、最も新しい情報のみを含み、そして1つのトラック書込み動作のみが、第2のリモートサイト23中のR2論理ボリューム40を更新するために必要になるであろう。そのような反復する書込み動作が単一のトラックに発生するとき、通信リンク36のためのバンド幅の要件を劇的に減少させることができる。要件を十分に減らすことができ、そして、T1ラインのような、少なくとも次のより低いレベルの通信バンド幅を使用することができることは、一般的にわかる。多くのアプリケーションでは、インターネットを通信リンク36として使用することができるようなレベルに更にバンド幅の要件を減少させることができる。

【0053】

DIFFERENTIAL SPLIT及びREESTABLISHコマンドの順番は、冗長なコピーのためのリポジトリとして、第2のリモートサイト23を使用することが望まれる限り、反復して継続することができる。一旦、冗長なコピーを維持する必要がなくなったら、ホストは、所定のCLOSE SESSIONコマンドを生成することができる。このカスケードコマンドは、レジスタ80中のPBビット位置の割り当てを開放し、カスケード動作を完了させるためにセッションフラグ81を消去する。

【0054】

本発明は、ローカルのプロダクションサイト21が、プロダクションデータ記憶設備中でデータを周期的に変更するための装置を表わすような好適な実施の形態に関して説明してきた。本発明によると、第1のリモートサイトは、プロダクションサイトでデータをミラリングするための第1のデータ記憶装置を含む。BCV/R1論理ボリューム34の形態の第2のデータ記憶装置は、2つの動作モードを有する。1つでは、BCV/R1論理ボリューム34は、R2論理ボリューム32のためのミラーとして働く。DIFFERENTIAL SPLIT又は同等のコマンドを発行することによって開始される他のモードでは、BCV/R1論理ボリューム34は、R2論理ボリューム32から分離され、BCV/R1トラックステータステーブル77中で識別されたデータをR2論理ボリューム40、又はBCV/R1論理ボリューム34からのデータのためのレシーバとして働く第2のリモートサイト23中の他の手段に伝達する。

【0055】

プロテクションビットテーブル75によって表わされるデータ変更記録装置は、第2の動作モードの間中、BCV/R1論理ボリューム34の内容がR2論理ボリューム40に転送されている間に、R2論理ボリューム32へのどのような変更も記録する。制御がBCV/R1論理ボリューム34を逆に第1の動作モードにシフトしているとき、データの変更の記録は、それらのデータブロック又はR2論理ボリューム32からBCV/R1論理ボリューム34に転送しなければならない他のデータブロックを識別する。この変更されたデータの記録は、システムが再びBCV/R1論理ボリューム34の動作を第2の動作モードにシフトさせるとき、BCV/R1論理ボリューム34から第2のリモートサイト23のようなデータ受信サイトに転送されるようなトラックを更に制御する。

【0056】

図1から6に示された具体的に開示された実施形態に行うことができる多くの変形がある。図1は、例えば、テーブル及びモジュールの特定の編成を表現する。第1のリモートサイト22の全体の動作から減じることなく、その情報の異なるコンフィギュレーションで置きかえることができよう。図1から6は、単一の論理ボリュームに関連する動作を更に表現する。一般的に、ネットワークは、図1から6に示される動作及び装置を繰り返すことによって、それぞれをこれと同じ方法で扱うことができる複数の論理ボリュームを含むであろう。更に、種々のモジュールの説明は、説明の目的のために特定の順番を確立するが；他の補助的な動作のために必要なような他の動作のいくつかの削除又は追加も含むことができる、他の順番でもよいであろう。そのような順番は、動作モードの間のそれぞれの転送をシステムのオペレータが開始するために、マニュアルで制御することができる。他には、カスケードセッションのすべて又は部分を自動制御とすることもできよう。このように、本発明は特定の実施形態に関して開示してきたが、本発明の真の精神及び範囲内の、すべての変形及び変更を包含することが、添付の請求項の目的である。

【図面の簡単な説明】

【図1】

プロダクション設備及び2つの地理的に遠隔の設備を含むデータ処理ネットワ

ークのブロック図である。

【図2】

プロダクション設備からリモート記憶設備へのコマンドの転送を示すフロー図である。

【図3】

DIFFERENTIAL SPLITコマンドに応答した図1のネットワークの動作を示す図である。

【図4】

図3で示された動作で使用する手順の詳細を示す図である。

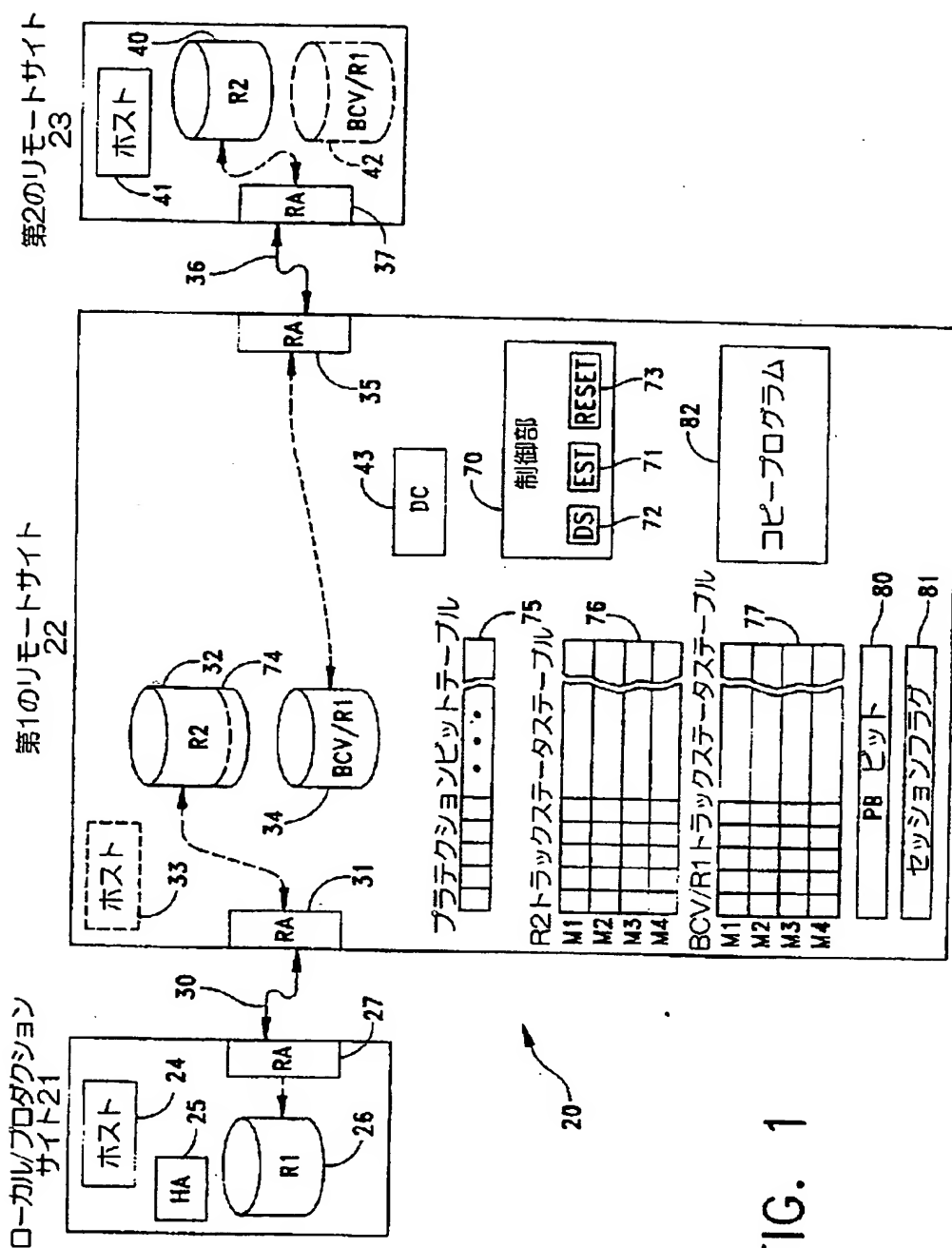
【図5】

RE-ESTABLISHコマンドに応答した図1のネットワークの動作を示す図である。

【図6】

DIFFERENTIAL SPLITコマンドを処理した後に起きる他の動作を示す図である。

【図1】



【図2】

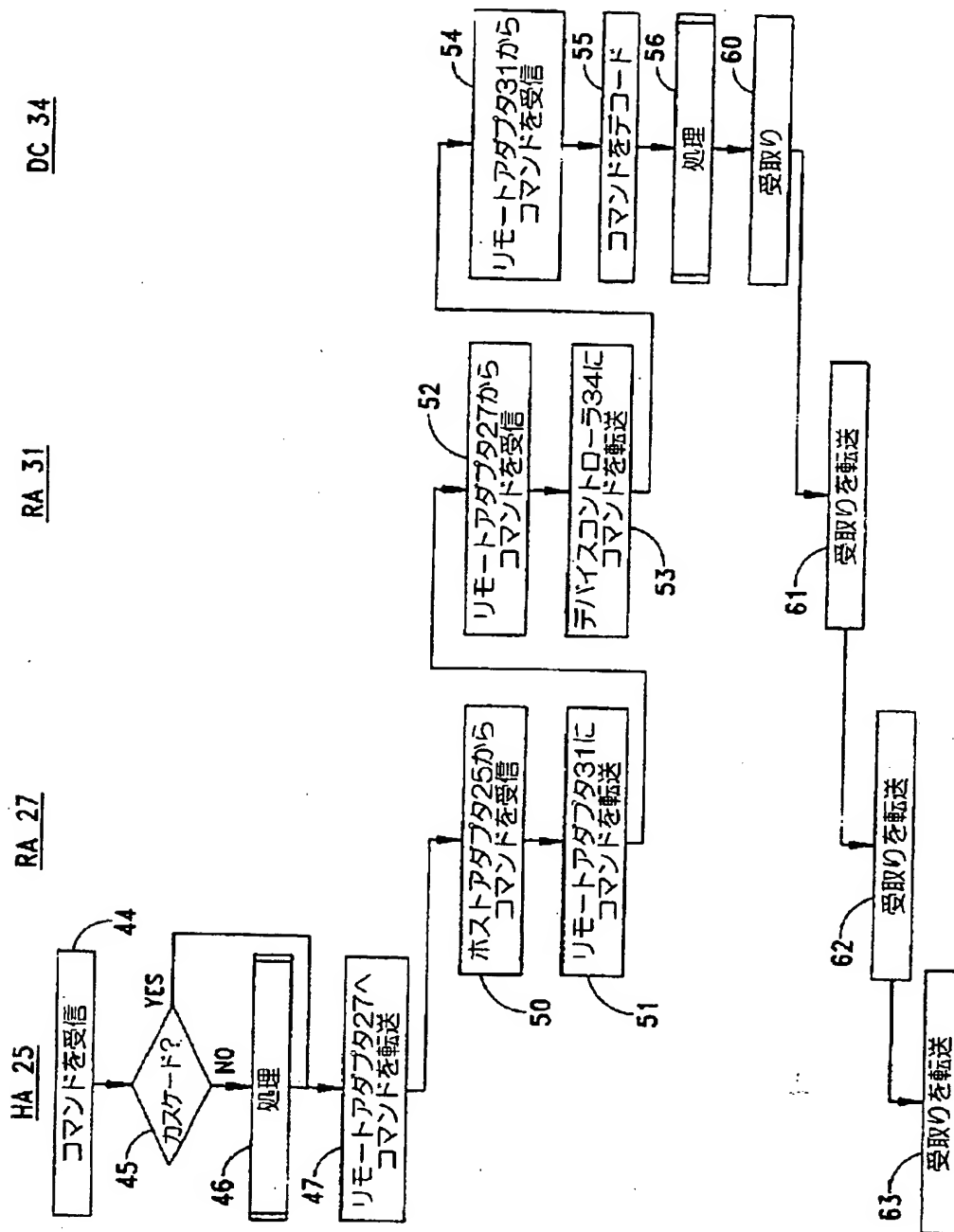


FIG. 2

【図3】

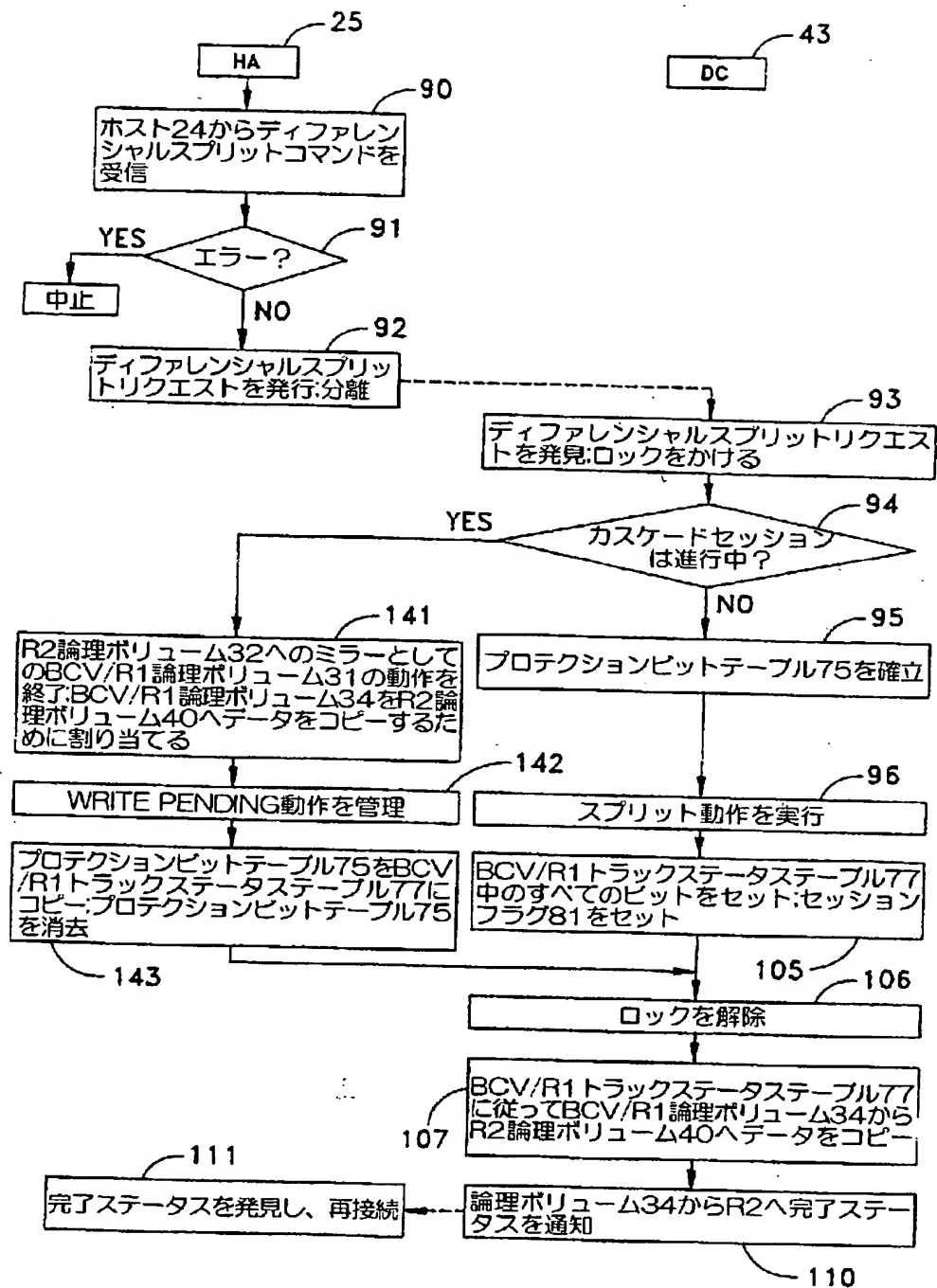


FIG. 3

【図4】

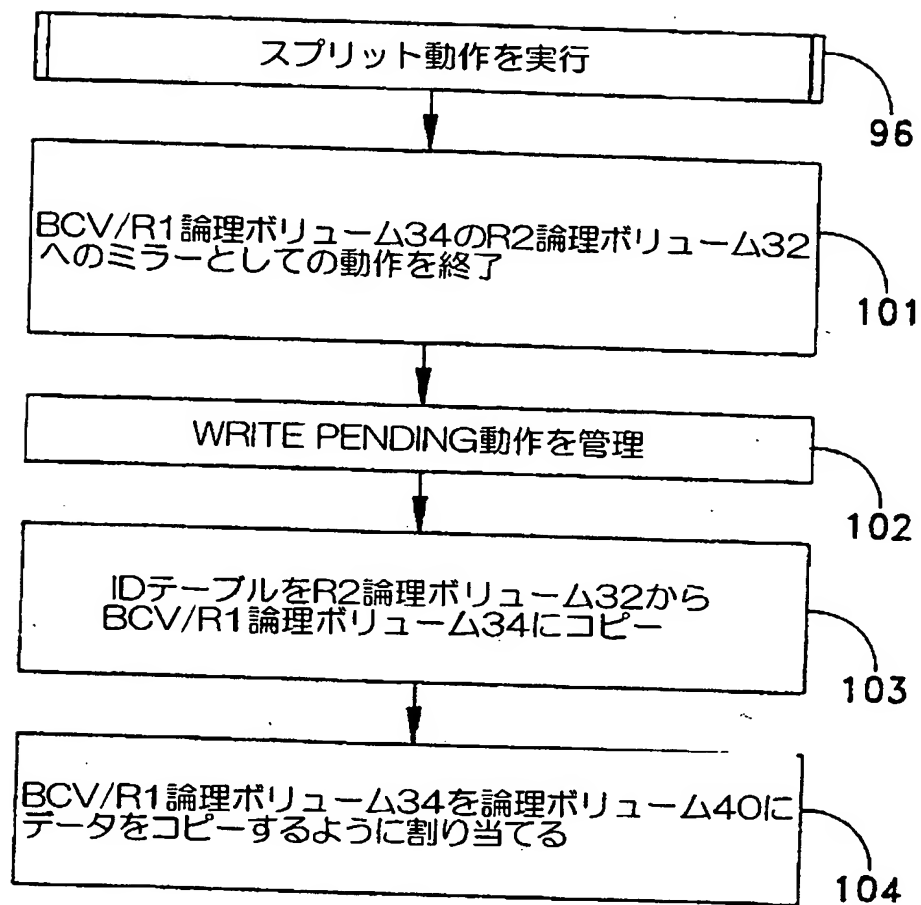


FIG. 4

【図5】

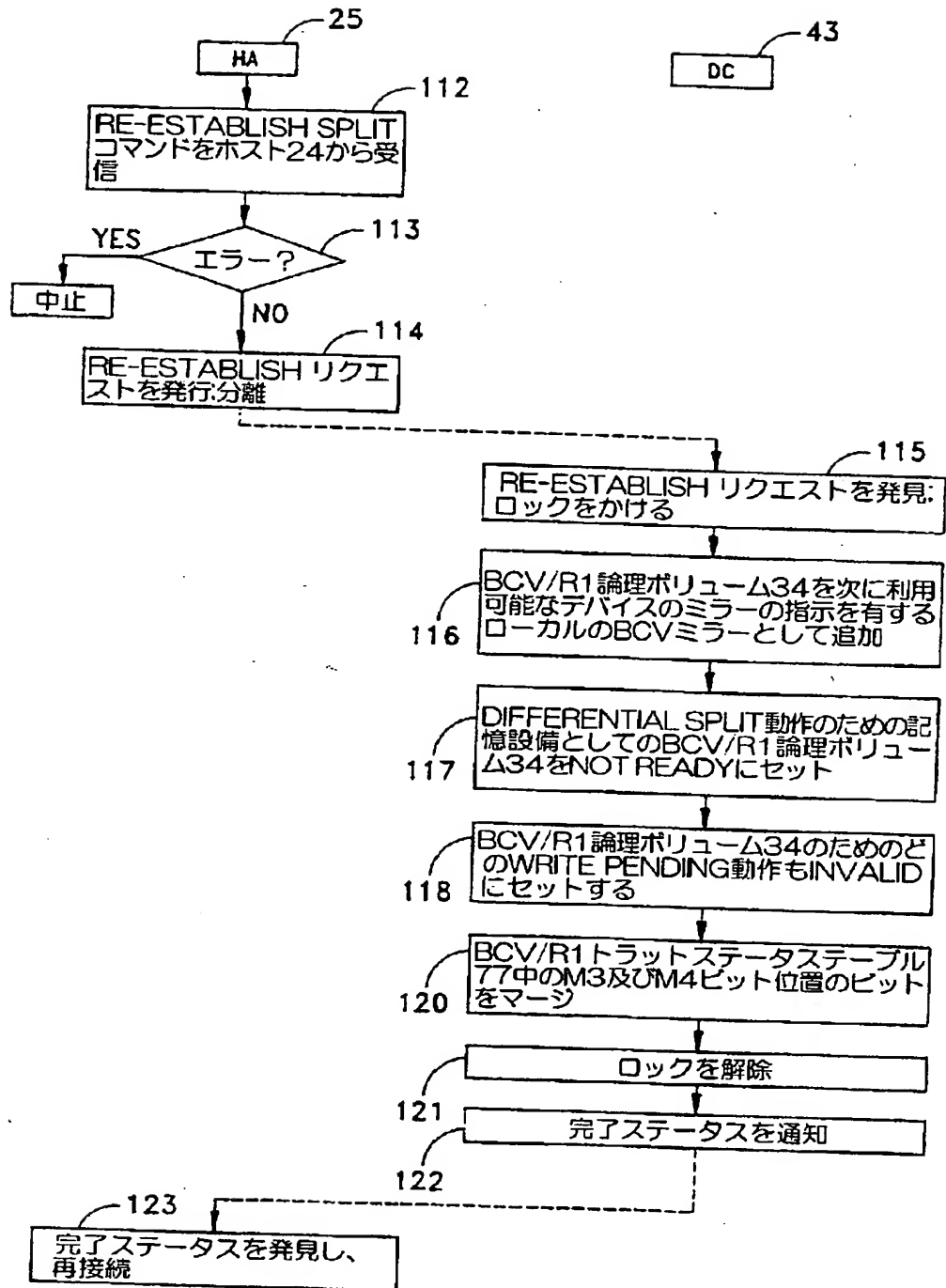


FIG. 5

【図6】

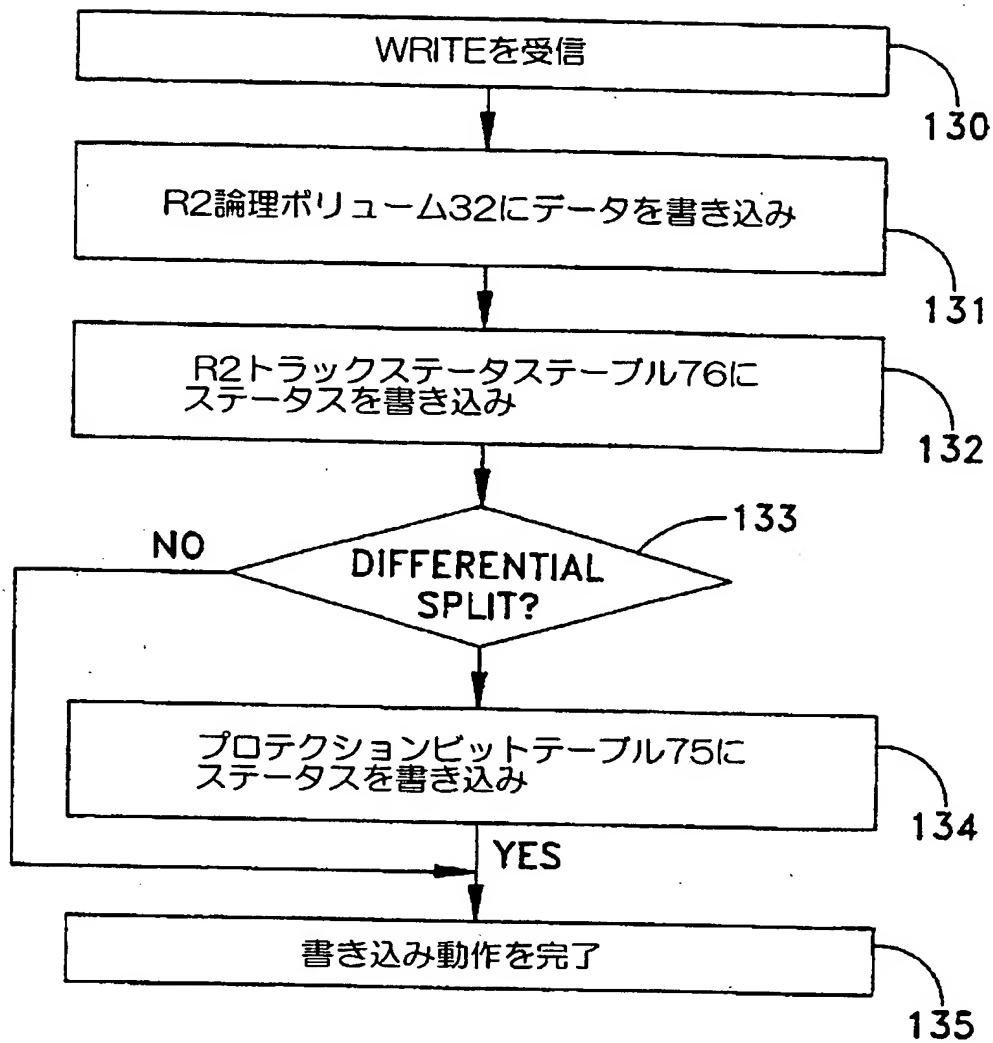


FIG. 6

【手続補正書】

【提出日】平成13年3月7日(2001.3.7)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】 データ処理システム中で、データを変更する手段と、データを受信する手段と、を有するデータネットワークであって、

A) データを受信するために前記データを変更する手段へ接続するための第1のデータ記憶装置と、

B) 第2のデータ記憶装置と、

C) 前記第1のデータ記憶装置中で前記データを変更する手段が行う変更の位置を識別するための第1及び第2の変更記録手段と、

D) 前記第1の変更記録手段中に記録された変更に従って、その間に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信するような第1の動作モードを確立するための第1の動作手段と、

E) 第2の動作モードが確立されているときに前記第1のデータ記憶装置中の記録された変更を反映している前記第2の変更記録手段中に記録された変更に従って、前記第2のデータ記憶装置から前記受信する手段にデータをコピーするための前記第2の動作モードを確立するための第2の動作手段と、を有することを特徴とするデータネットワーク。

【請求項2】 前記第1及び第2の変更記録手段は、それぞれ第1及び第2のテーブルを含み、前記データネットワーク内の転送は、データブロックとして起こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の情報のための記憶装置を含み、前記第2の変更記録手段中の前記第2のテーブルは、当該別の情報を受信し、前記第2の動作手段は、前記第2のテーブルの内容を、前記第2のデータ記憶装置から前記データを受信する手段へのデータの転

送を制御する第3のテーブルに転送することを特徴とする請求項1に記載のデータネットワーク。

【請求項3】 前記データ記憶設備内の転送は、データブロックとして起こり、前記第1のデータ記憶装置の一部は、それぞれのデータブロックへの変更についての情報を含み、前記データ記憶設備は、第3の変更記録手段を更に含み、及び、前記第2の動作手段は、第2のモードが確立されているときに、前記第1のデータ記憶装置の一部の内容を前記第3の変更記録手段に転送することを特徴とする請求項1に記載のデータネットワーク。

【請求項4】 その間に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可能にするようなカスケードセッションを実行するための、前記変更する手段から反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に含むことを特徴とする請求項3に記載のデータネットワーク。

【請求項5】 前記第1、第2及び第3の変更記録手段は、それぞれ第1、第2及び第3のテーブルを含み、前記第2のテーブルは、記憶された別の情報から変更情報を受信し、前記ネットワークは、データを転送するためのコピー手段を更に含み、前記第2の動作手段は、前記コピー手段が前記第2のデータ記憶装置中のすべてのデータを前記受信する手段に転送することができるようにするためのカスケードセッションの第1の反復の間に、第2のテーブルを初期化し、かつ、前記第3のテーブル中のすべてのデータを第1の状態に設定することを特徴とする請求項4に記載のデータネットワーク。

【請求項6】 データを変更する設備と、データを受信する設備の間でデータを転送するための方法であって、

- A) 変更されたデータを受信するための第1のデータ記憶装置を確立するステップと、
- B) 第2のデータ記憶装置を確立するステップと、
- C) 第1及び第2の位置に、データを変更するための設備によって作り出される前記第1のデータ記憶装置中の変更の識別を記録するステップと、
- D) 前記第1の位置中の識別された変更に従って、その間に前記第2のデータ記

憶装置が前記第1のデータ記憶装置からデータを受信するような第1の動作モードを確立するステップと、

E) 第2の動作モードが確立されているときに識別される第2の位置中の変更に従って、前記第2のデータ記憶装置から前記受信する設備にデータをコピーするための第2の動作モードを確立するステップと、を有することを特徴とする方法。

【請求項7】 データ転送は、データブロックとして起こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の情報のための記憶装置を含み、変更を識別する前記ステップは、変更の識別を記録するための第3の位置を確立することを含み、前記第2の動作モードの確立は、前記受信する設備へのコピーができるようにするために、第2の位置の内容を第3の位置へ転送することを特徴とする請求項6に記載の方法。

【請求項8】 データ転送は、データブロックとして起こり、前記第1のデータ記憶装置の一部は、それぞれのデータブロックへの変更についての情報を含み、前記第1のデータ記憶装置の一部の内容は、第2の動作モードが確立されているときに、最初の識別情報を提供するために前記第2の位置に転送されることを特徴とする請求項6に記載の方法。

【請求項9】 前記変更する設備からの第1及び第2のコマンドの反復した生成は、その間に前記第1及び第2の動作モードが相互に排他的に使用可能にされるようなカスケードセッションを反復して実行することを特徴とする請求項8に記載の方法。

【請求項10】 データは、コピーによって転送され、前記第2の動作モードの最初の動作は、当該コピーが前記第2のデータ記憶装置中のすべてのデータを前記受信する設備に転送するようなカスケードセッションにおける最初の反復の間に、前記第2の位置を初期化し、及び、第3の位置中のすべてのデータを第1の状態に設定することを特徴とする請求項9に記載の方法。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 99/23939**A. CLASSIFICATION OF SUBJECT MATTER**
IPC 7 G06F11/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 45790 A (EMC CORPORATION) 4 December 1997 (1997-12-04) page 54, line 15 -page 55, line 15; figure 22	1,8,20, 27
A	WO 98 20419 A (VINCA CORPORATION) 14 May 1998 (1998-05-14) page 5, line 15 -page 7, line 34	1,8,20, 27
A	US 5 799 147 A (SHANNON) 25 August 1998 (1998-08-25) the whole document	1,8,20, 27
A	US 5 857 208 A (OFEX) 5 January 1999 (1999-01-05) abstract	1,8,20, 27
	-/-	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (see specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"a" document member of the same patent family

Date of the actual completion of the international search

23 March 2000

Date of mailing of the international search report

30/03/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentamt 2
NL - 2200 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 681 epo nl,
Fax (+31-70) 340-8016

Authorized officer

Absalom, R

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Internat. J. Application No.
PCT/US 99/23939

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 86 01018 A (AMERICAN TELEPHONE AND TELEGRAPH COMPANY) <u>13 February 1986 (1986-02-13)</u>	
A	US 5 634 052 A (MORRIS) <u>27 May 1997 (1997-05-27)</u>	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.
PCT/US 99/23939

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9745790	A	04-12-1997	AU 3224897 A EP 0902923 A	05-01-1998 24-03-1999
WO 9820419	A	14-05-1998	US 5835953 A AU 5174098 A	10-11-1998 29-05-1998
US 5799147	A	25-08-1998	US 5852713 A	22-12-1998
US 5857208	A	05-01-1999	NONE	
WO 8601018	A	13-02-1986	US 4686620 A CA 1237201 A EP 0191036 A JP 6001447 B JP 61502846 T KR 9308684 B	11-08-1987 24-05-1988 20-08-1986 05-01-1994 04-12-1986 11-09-1993
US 5634052	A	27-05-1997	NONE	

フロントページの続き

Fターム(参考) 5B018 GA04 HA05 MA12 QA15
5B082 DE03 GB02
5B083 AA05 CC04 CD11 EE08